

OIPE

RAW SEQUENCE LISTING

DATE: 12/07/2001

PATENT APPLICATION: US/09/991,053

TIME: 12:15:46

Input Set : N:\Crf3\RULE60\09991053.txt

Output Set: N:\CRF3\12072001\I991053.raw

3 <110> APPLICANT: Harrison, Stephen
 4 Chen, Xiaojiang
 6 <120> TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
 7 virus
 9 <130> FILE REFERENCE: Harvard/Harrison 12687/1120
 11 <140> CURRENT APPLICATION NUMBER: 09/991,053
 12 <141> CURRENT FILING DATE: 2001-11-21
 14 <150> PRIOR APPLICATION NUMBER: 09/520,781
 15 <151> PRIOR FILING DATE: 2000-03-18
 17 <150> PRIOR APPLICATION NUMBER: 60/125208
 18 <151> PRIOR FILING DATE: 1999-03-18
 20 <150> PRIOR APPLICATION NUMBER: 60/148544
 21 <151> PRIOR FILING DATE: 1999-08-12
 23 <160> NUMBER OF SEQ ID NOS: 4
 25 <170> SOFTWARE: PatentIn Ver. 2.1
 27 <210> SEQ ID NO: 1
 28 <211> LENGTH: 531
 29 <212> TYPE: PRT
 30 <213> ORGANISM: Human papillomavirus type 16
 32 <400> SEQUENCE: 1
 33 Met Gln Val Thr Phe Ile Tyr Ile Leu Val Ile Thr Cys Tyr Glu Asn
 34 1 5 10 15
 36 Asp Val Asn Val Tyr His Ile Phe Phe Gln Met Ser Leu Trp Leu Pro
 37 20 25 30
 39 Ser Glu Ala Thr Val Tyr Leu Pro Pro Val Pro Val Ser Lys Val Val
 40 35 40 45
 42 Ser Thr Asp Glu Tyr Val Ala Arg Thr Asn Ile Tyr Tyr His Ala Gly
 43 50 55 60
 45 Thr Ser Arg Leu Leu Ala Val Gly His Pro Tyr Phe Pro Ile Lys Lys
 46 65 70 75 80
 48 Pro Asn Asn Asn Lys Ile Leu Val Pro Lys Val Ser Gly Leu Gln Tyr
 49 85 90 95
 51 Arg Val Phe Arg Ile His Leu Pro Asp Pro Asn Lys Phe Gly Phe Pro
 52 100 105 110
 54 Asp Thr Ser Phe Tyr Asn Pro Asp Thr Gln Arg Leu Val Trp Ala Cys
 55 115 120 125
 57 Val Gly Val Glu Val Gly Arg Gly Gln Pro Leu Gly Val Gly Ile Ser
 58 130 135 140
 60 Gly His Pro Leu Leu Asn Lys Leu Asp Asp Thr Glu Asn Ala Ser Ala
 61 145 150 155 160
 63 Tyr Ala Ala Asn Ala Gly Val Asp Asn Arg Glu Cys Ile Ser Met Asp
 64 165 170 175
 66 Tyr Lys Gln Thr Gln Leu Cys Leu Ile Gly Cys Lys Pro Pro Ile Gly
 67 180 185 190
 69 Glu His Trp Gly Lys Gly Ser Pro Cys Thr Asn Val Ala Val Asn Pro
 70 195 200 205
 72 Gly Asp Cys Pro Pro Leu Glu Leu Ile Asn Thr Val Ile Gln Asp Gly

ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/991,053

DATE: 12/07/2001

TIME: 12:15:46

Input Set : N:\Crf3\RULE60\09991053.txt

Output Set: N:\CRF3\12072001\I991053.raw

```

73      210      215      220
75 Asp Met Val His Thr Gly Phe Gly Ala Met Asp Phe Thr Thr Leu Gln
76 225      230      235      240
78 Ala Asn Lys Ser Glu Val Pro Leu Asp Ile Cys Thr Ser Ile Cys Lys
79      245      250      255
81 Tyr Pro Asp Tyr Ile Lys Met Val Ser Glu Pro Tyr Gly Asp Ser Leu
82      260      265      270
84 Phe Phe Tyr Leu Arg Arg Glu Gln Met Phe Val Arg His Leu Phe Asn
85      275      280      285
87 Arg Ala Gly Thr Val Gly Glu Asn Val Pro Asp Asp Leu Tyr Ile Lys
88      290      295      300
90 Gly Ser Gly Ser Thr Ala Asn Leu Ala Ser Ser Asn Tyr Phe Pro Thr
91 305      310      315      320
93 Pro Ser Gly Ser Met Val Thr Ser Asp Ala Gln Ile Phe Asn Lys Pro
94      325      330      335
96 Tyr Trp Leu Gln Arg Ala Gln Gly His Asn Asn Gly Ile Cys Trp Gly
97      340      345      350
99 Asn Gln Leu Phe Val Thr Val Val Asp Thr Thr Arg Ser Thr Asn Met
100      355      360      365
102 Ser Leu Cys Ala Ala Ile Ser Thr Ser Glu Thr Thr Tyr Lys Asn Thr
103      370      375      380
105 Asn Phe Lys Glu Tyr Leu Arg His Gly Glu Glu Tyr Asp Leu Gln Phe
106 385      390      395      400
108 Ile Phe Gln Leu Cys Lys Ile Thr Leu Thr Ala Asp Val Met Thr Tyr
109      405      410      415
111 Ile His Ser Met Asn Ser Thr Ile Leu Glu Asp Trp Asn Phe Gly Leu
112      420      425      430
114 Gln Pro Pro Gly Gly Thr Leu Glu Asp Thr Tyr Arg Phe Val Thr
115      435      440      445
117 Gln Ala Ile Ala Cys Gln Lys His Thr Pro Pro Ala Pro Lys Glu Asp
118      450      455      460
120 Asp Pro Leu Lys Lys Tyr Thr Phe Trp Glu Val Asn Leu Lys Glu Lys
121 465      470      475      480
123 Phe Ser Ala Asp Leu Asp Gln Phe Pro Leu Gly Arg Lys Phe Leu Leu
124      485      490      495
126 Gln Ala Gly Leu Lys Ala Lys Pro Lys Phe Thr Leu Gly Lys Arg Lys
127      500      505      510
129 Ala Thr Pro Thr Thr Ser Ser Thr Ser Thr Thr Ala Lys Arg Lys Lys
130      515      520      525
132 Arg Lys Leu
133      530
136 <210> SEQ ID NO: 2
137 <211> LENGTH: 473
138 <212> TYPE: PRT
139 <213> ORGANISM: Human papillomavirus type 16
141 <400> SEQUENCE: 2
142 Met Arg His Lys Arg Ser Ala Lys Arg Thr Lys Arg Ala Ser Ala Thr
143 1      5      10      15
145 Gln Leu Tyr Lys Thr Cys Lys Gln Ala Gly Thr Cys Pro Pro Asp Ile

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/991,053

DATE: 12/07/2001

TIME: 12:15:46

Input Set : N:\Crf3\RULE60\09991053.txt

Output Set: N:\CRF3\12072001\I991053.raw

```

146          20          25          30
148 Ile Pro Lys Val Glu Gly Lys Thr Ile Ala Glu Gln Ile Leu Gln Tyr
149          35          40          45
151 Gly Ser Met Gly Val Phe Phe Gly Gly Leu Gly Ile Gly Thr Gly Ser
152          50          55          60
154 Gly Thr Gly Gly Arg Thr Gly Tyr Ile Pro Leu Gly Thr Arg Pro Pro
155 65          70          75          80
157 Thr Ala Thr Asp Thr Leu Ala Pro Val Arg Pro Pro Leu Thr Val Asp
158          85          90          95
160 Pro Val Gly Pro Ser Asp Pro Ser Ile Val Ser Leu Val Glu Thr
161          100         105         110
163 Ser Phe Ile Asp Ala Gly Ala Pro Thr Ser Val Pro Ser Ile Pro Pro
164          115         120         125
166 Asp Val Ser Gly Phe Ser Ile Thr Thr Ser Thr Asp Thr Thr Pro Ala
167          130         135         140
169 Ile Leu Asp Ile Asn Asn Thr Val Thr Thr Val Thr Thr His Asn Asn
170 145         150         155         160
172 Pro Thr Phe Thr Asp Pro Ser Val Leu Gln Pro Pro Thr Pro Ala Glu
173         165         170         175
175 Thr Gly Gly His Phe Thr Leu Ser Ser Ser Thr Ile Ser Thr His Asn
176         180         185         190
178 Tyr Glu Glu Ile Pro Met Asp Thr Phe Ile Val Ser Thr Asn Pro Asn
179         195         200         205
181 Thr Val Thr Ser Ser Thr Pro Ile Pro Gly Ser Arg Pro Val Ala Arg
182         210         215         220
184 Leu Gly Leu Tyr Ser Arg Thr Thr Gln Gln Val Lys Val Val Asp Pro
185 225         230         235         240
187 Ala Phe Val Thr Thr Pro Thr Lys Leu Ile Thr Tyr Asp Asn Pro Ala
188         245         250         255
190 Tyr Glu Gly Ile Asp Val Asp Asn Thr Leu Tyr Phe Ser Ser Asn Asp
191         260         265         270
193 Asn Ser Ile Asn Ile Ala Pro Asp Pro Asp Phe Leu Asp Ile Val Ala
194         275         280         285
196 Leu His Arg Pro Ala Leu Thr Ser Arg Arg Thr Gly Ile Arg Tyr Ser
197         290         295         300
199 Arg Ile Gly Asn Lys Gln Thr Leu Arg Thr Arg Ser Gly Lys Ser Ile
200 305         310         315         320
202 Gly Ala Lys Val His Tyr Tyr Tyr Asp Leu Ser Thr Ile Asp Pro Ala
203         325         330         335
205 Glu Glu Ile Glu Leu Gln Thr Ile Thr Pro Ser Thr Tyr Thr Thr Thr
206         340         345         350
208 Ser His Ala Ala Ser Pro Thr Ser Ile Asn Asn Gly Leu Tyr Asp Ile
209         355         360         365
211 Tyr Ala Asp Asp Phe Ile Thr Asp Thr Ser Thr Thr Pro Val Pro Ser
212         370         375         380
214 Val Pro Ser Thr Ser Leu Ser Gly Tyr Ile Pro Ala Asn Thr Thr Ile
215 385         390         395         400
217 Pro Phe Gly Gly Ala Tyr Asn Ile Pro Leu Val Ser Gly Pro Asp Ile
218         405         410         415

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/991,053

DATE: 12/07/2001

TIME: 12:15:46

Input Set : N:\Crf3\RULE60\09991053.txt

Output Set: N:\CRF3\12072001\I991053.raw

```

220 Pro Ile Asn Ile Thr Asp Gln Ala Pro Ser Leu Ile Pro Ile Val Pro
221           420           425           430
223 Gly Ser Pro Gln Tyr Thr Ile Ile Ala Asp Ala Gly Asp Phe Tyr Leu
224           435           440           445
226 His Pro Ser Tyr Tyr Met Leu Arg Lys Arg Arg Lys Arg Leu Pro Tyr
227           450           455           460
229 Phe Phe Ser Asp Val Ser Leu Ala Ala
230 465           470
233 <210> SEQ ID NO: 3
234 <211> LENGTH: 568
235 <212> TYPE: PRT
236 <213> ORGANISM: Human papillomavirus type 18
238 <400> SEQUENCE: 3
239 Met Cys Leu Tyr Thr Arg Val Leu Ile Leu His Tyr His Leu Leu Pro
240   1           5           10           15
242 Leu Tyr Gly Pro Leu Tyr His Pro Arg Pro Leu Pro Leu His Ser Ile
243           20           25           30
245 Leu Val Tyr Met Val His Ile Ile Ile Cys Gly His Tyr Ile Ile Leu
246           35           40           45
248 Phe Leu Arg Asn Val Asn Val Phe Pro Ile Phe Leu Gln Met Ala Leu
249           50           55           60
251 Trp Arg Pro Ser Asp Asn Thr Val Tyr Leu Pro Pro Pro Ser Val Ala
252   65           70           75           80
254 Arg Val Val Asn Thr Asp Asp Tyr Val Thr Pro Thr Ser Ile Phe Tyr
255           85           90           95
257 His Ala Gly Ser Ser Arg Leu Leu Thr Val Gly Asn Pro Tyr Phe Arg
258           100          105          110
260 Val Pro Ala Gly Gly Gly Asn Lys Gln Asp Ile Pro Lys Val Ser Ala
261           115          120          125
263 Tyr Gln Tyr Arg Val Phe Arg Val Gln Leu Pro Asp Pro Asn Lys Phe
264           130          135          140
266 Gly Leu Pro Asp Thr Ser Ile Tyr Asn Pro Glu Thr Gln Arg Leu Val
267 145           150          155          160
269 Trp Ala Cys Ala Gly Val Glu Ile Gly Arg Gly Gln Pro Leu Gly Val
270           165          170          175
272 Gly Leu Ser Gly His Pro Phe Tyr Asn Lys Leu Asp Asp Thr Glu Ser
273           180          185          190
275 Ser His Ala Ala Thr Ser Asn Val Ser Glu Asp Val Arg Asp Asn Val
276           195          200          205
278 Ser Val Asp Tyr Lys Gln Thr Gln Leu Cys Ile Leu Gly Cys Ala Pro
279           210          215          220
281 Ala Ile Gly Glu His Trp Ala Lys Gly Thr Ala Cys Lys Ser Arg Pro
282 225           230          235          240
284 Leu Ser Gln Gly Asp Cys Pro Pro Leu Glu Leu Lys Asn Thr Val Leu
285           245          250          255
287 Glu Asp Gly Asp Met Val Asp Thr Gly Tyr Gly Ala Met Asp Phe Ser
288           260          265          270
290 Thr Leu Gln Asp Thr Lys Cys Glu Val Pro Leu Asp Ile Cys Gln Ser
291           275          280          285

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/991,053

DATE: 12/07/2001

TIME: 12:15:46

Input Set : N:\Crif3\RULE60\09991053.txt

Output Set: N:\CRF3\12072001\I991053.raw

```

293 Ile Cys Lys Tyr Pro Asp Tyr Leu Gln Met Ser Ala Asp Pro Tyr Gly
294      290      295      300
296 Asp Ser Met Phe Phe Cys Leu Arg Arg Glu Gln Leu Phe Ala Arg His
297 305      310      315      320
299 Phe Trp Asn Arg Ala Gly Thr Met Gly Asp Thr Val Pro Gln Ser Leu
300      325      330      335
302 Tyr Ile Lys Gly Thr Gly Met Pro Ala Ser Pro Gly Ser Cys Val Tyr
303      340      345      350
305 Ser Pro Ser Pro Ser Gly Ser Ile Val Thr Ser Asp Ser Gln Leu Phe
306      355      360      365
308 Asn Lys Pro Tyr Trp Leu His Lys Ala Gln Gly His Asn Asn Gly Val
309      370      375      380
311 Cys Trp His Asn Gln Leu Phe Val Thr Val Val Asp Thr Thr Pro Ser
312 385      390      395      400
314 Thr Asn Leu Thr Ile Cys Ala Ser Thr Gln Ser Pro Val Pro Gly Gln
315      405      410      415
317 Tyr Asp Ala Thr Lys Phe Lys Gln Tyr Ser Arg His Val Glu Glu Tyr
318      420      425      430
320 Asp Leu Gln Phe Ile Phe Gln Leu Cys Thr Ile Thr Leu Thr Ala Asp
321      435      440      445
323 Val Met Ser Tyr Ile His Ser Met Asn Ser Ser Ile Leu Glu Asp Trp
324      450      455      460
326 Asn Phe Gly Val Pro Pro Pro Thr Thr Ser Leu Val Asp Thr Tyr
327 465      470      475      480
329 Arg Phe Val Gln Ser Val Ala Ile Thr Cys Gln Lys Asp Ala Ala Pro
330      485      490      495
332 Ala Glu Asn Lys Asp Pro Tyr Asp Lys Leu Lys Phe Trp Asn Val Asp
333      500      505      510
335 Leu Lys Glu Lys Phe Ser Leu Asp Leu Asp Gln Tyr Pro Leu Gly Arg
336      515      520      525
338 Lys Phe Leu Val Gln Ala Gly Leu Arg Arg Lys Pro Thr Ile Gly Pro
339      530      535      540
341 Arg Lys Arg Ser Ala Pro Ser Ala Thr Thr Ser Ser Lys Pro Ala Lys
342 545      550      555      560
344 Arg Val Arg Val Arg Ala Arg Lys
345      565
348 <210> SEQ ID NO: 4
349 <211> LENGTH: 462
350 <212> TYPE: PRT
351 <213> ORGANISM: Human papillomavirus type 18
353 <400> SEQUENCE: 4
354 Met Val Ser His Arg Ala Ala Arg Arg Lys Arg Ala Ser Val Thr Asp
355 1      5      10      15
357 Leu Tyr Lys Thr Cys Lys Gln Ser Gly Thr Cys Pro Pro Asp Val Val
358      20      25      30
360 Pro Lys Val Glu Gly Thr Thr Leu Ala Asp Lys Ile Leu Gln Trp Ser
361      35      40      45
363 Ser Leu Gly Ile Phe Leu Gly Gly Leu Gly Ile Gly Thr Gly Ser Gly
364      50      55      60

```

VERIFICATION SUMMARY

DATE: 12/07/2001

PATENT APPLICATION: US/09/991,053

TIME: 12:15:47

Input Set : N:\Crf3\RULE60\09991053.txt

Output Set: N:\CRF3\12072001\I991053.raw